



ENTERED PCT10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/088,027

DATE: 04/04/2002

TIME: 15:39:20

Input Set : A:\seq list.txt

Output Set: N:\CRF3\04042002\J088027.raw

```

3 <110> APPLICANT: BASF Aktiengesellschaft
5 <120> TITLE OF INVENTION: GMP synthetase from plants
7 <130> FILE REFERENCE: DE 19947490.7
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/088,027
C--> 10 <141> CURRENT FILING DATE: 2002-03-14
12 <160> NUMBER OF SEQ ID NOS: 4
14 <170> SOFTWARE: PatentIn Vers. 2.0
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1973
18 <212> TYPE: DNA
19 <213> ORGANISM: Nicotiana tabacum
21 <220> FEATURE:
22 <221> NAME/KEY: CDS
23 <222> LOCATION: (65)..(1678)
25 <400> SEQUENCE: 1
26 gaattcgga cgagatttct ctctatcttt cttcctccca cccaccaccc accctcccct 60
28 agca atg gaa cct caa aca cag gcg aag aaa tca aac ctc gta cta atc 109
29 Met Glu Pro Gln Thr Gln Ala Lys Lys Ser Asn Leu Val Leu Ile
30 1 5 10 15
32 cta gac tac ggt tct cag tac act cac cta atc acc cgc cga atc cga 157
33 Leu Asp Tyr Gly Ser Gln Tyr Thr His Leu Ile Thr Arg Arg Ile Arg
34 20 25 30
36 agc cta tca att ttc tca ctc acc att aac ggc acc tct tcg tta gac 205
37 Ser Leu Ser Ile Phe Ser Leu Thr Ile Asn Gly Thr Ser Ser Leu Asp
38 35 40 45
40 tcc ata aaa gaa ctc gac cca cgt gtc att atc ctc tcg ggt gga ccc 253
41 Ser Ile Lys Glu Leu Asp Pro Arg Val Ile Ile Leu Ser Gly Gly Pro
42 50 55 60
44 cac agc gtc cac gct gac ggc gca ccg tgt ttc cca cct ggg ttc atc 301
45 His Ser Val His Ala Asp Gly Ala Pro Cys Phe Pro Pro Gly Phe Ile
46 65 70 75
48 gaa tac gtc gag tca cgt ggg att cac gtg ttg ggt ata tgt tat ggg 349
49 Glu Tyr Val Glu Ser Arg Gly Ile His Val Leu Gly Ile Cys Tyr Gly
50 80 85 90 95
52 ctg cag ttg att gtt cag aaa ctt ggc ggg gtt gtg aaa att gga gag 397
53 Leu Gln Leu Ile Val Gln Lys Leu Gly Gly Val Val Lys Ile Gly Glu
54 100 105 110
56 aaa cat gag tat ggg aga atg gaa att gag gtt gga aag aat gtt gtt 445
57 Lys His Glu Tyr Gly Arg Met Glu Ile Glu Val Gly Lys Asn Val Val
58 115 120 125
60 ggg ggg ttg ttt ggg aat acg gaa att ggt gat aaa cag gtg gtt tgg 493
61 Gly Gly Leu Phe Gly Asn Thr Glu Ile Gly Asp Lys Gln Val Val Trp
62 130 135 140

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/088,027

DATE: 04/04/2002

TIME: 15:39:20

Input Set : A:\seq list.txt

Output Set: N:\CRF3\04042002\J088027.raw

```

64  atg agc cac ggt gat gag gct gtg aaa ttg ccg gaa ggg ttt gag gtt 541
65  Met Ser His Gly Asp Glu Ala Val Lys Leu Pro Glu Gly Phe Glu Val
66      145      150      155
68  gtg gcg agg agt agt cag ggt gct gtt gct gct att gag aat cgg gaa 589
69  Val Ala Arg Ser Ser Gln Gly Ala Val Ala Ala Ile Glu Asn Arg Glu
70  160      165      170      175
72  cgg agg ttt tat ggg ctg cag tat cat ccc gag gta acg cac tcg act 637
73  Arg Arg Phe Tyr Gly Leu Gln Tyr His Pro Glu Val Thr His Ser Thr
74      180      185      190
76  gaa ggg atg aga aca tta aga cac ttt ctg ttt gat gta tgt ggc gtt 685
77  Glu Gly Met Arg Thr Leu Arg His Phe Leu Phe Asp Val Cys Gly Val
78      195      200      205
80  aca gct ggc tgg aag atg gaa gat gtt ctg gag gaa gaa ata aaa gtt 733
81  Thr Ala Gly Trp Lys Met Glu Asp Val Leu Glu Glu Glu Ile Lys Val
82      210      215      220
84  atc aaa ggt atg gtt gga cct gaa gat cac gtg att tgt gct tta tct 781
85  Ile Lys Gly Met Val Gly Pro Glu Asp His Val Ile Cys Ala Leu Ser
86      225      230      235
88  ggt ggt gtt gat tcc aca gtt gca gct aaa ttg gta cac aag gct atc 829
89  Gly Gly Val Asp Ser Thr Val Ala Ala Lys Leu Val His Lys Ala Ile
90  240      245      250      255
92  ggg gac agg ctt cac tgt gtt ttt gtt gat aat ggt cta tta agg tat 877
93  Gly Asp Arg Leu His Cys Val Phe Val Asp Asn Gly Leu Leu Arg Tyr
94      260      265      270
96  aag gag aga gaa agg gtg atg gaa ctc ttt gag aag cgc ctt cat ttg 925
97  Lys Glu Arg Glu Arg Val Met Glu Leu Phe Glu Lys Arg Leu His Leu
98      275      280      285
100 cct gtt acc tgt gtc gat gct aca gaa gaa ttt ctc agc aaa cta aaa 973
101 Pro Val Thr Cys Val Asp Ala Thr Glu Glu Phe Leu Ser Lys Leu Lys
102      290      295      300
104 ggc gta aca gaa cct gaa atg aag agg aaa ata att ggg aag gag ttc 1021
105 Gly Val Thr Glu Pro Glu Met Lys Arg Lys Ile Ile Gly Lys Glu Phe
106      305      310      315
108 atc aac ata ttt gat ctt ttt gcc cat gat gtg gag gaa aaa gta ggg 1069
109 Ile Asn Ile Phe Asp Leu Phe Ala His Asp Val Glu Glu Lys Val Gly
110 320      325      330      335
112 aaa aaa cct agt tac cta gtc caa gga acc ttg tat cct gat gta ata 1117
113 Lys Lys Pro Ser Tyr Leu Val Gln Gly Thr Leu Tyr Pro Asp Val Ile
114      340      345      350
116 gag tct tgt cct cca cct gga agt gga aga aca cat tct cat aca atc 1165
117 Glu Ser Cys Pro Pro Pro Gly Ser Gly Arg Thr His Ser His Thr Ile
118      355      360      365
120 aag agc cat cat aat gtt gga ggt ctt cca aag gac atg aag ctg aag 1213
121 Lys Ser His His Asn Val Gly Gly Leu Pro Lys Asp Met Lys Leu Lys
122      370      375      380
124 ctc atc gag cca ctg aaa ctt cta ttc aag gat gag gtt cgt gaa ttg 1261
125 Leu Ile Glu Pro Leu Lys Leu Leu Phe Lys Asp Glu Val Arg Glu Leu
126      385      390      395
128 gga aag att ttg gat ata tct gag gac ttt ctt aaa cgc cac ccg ttc 1309

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/088,027

DATE: 04/04/2002

TIME: 15:39:20

Input Set : A:\seq list.txt

Output Set: N:\CRF3\04042002\J088027.raw

```

129 Gly Lys Ile Leu Asp Ile Ser Glu Asp Phe Leu Lys Arg His Pro Phe
130 400                               405                               410                               415
132 cct ggg ccc gga ctc gct gtg cga att cca ggt gat gtc aca gca ggg 1357
133 Pro Gly Pro Gly Leu Ala Val Arg Ile Pro Gly Asp Val Thr Ala Gly
134                               420                               425                               430
136 aat tcc ttg gat att ctt cgt cag gtt gat gaa atc ttc att caa tca 1405
137 Asn Ser Leu Asp Ile Leu Arg Gln Val Asp Glu Ile Phe Ile Gln Ser
138                               435                               440                               445
140 atc aga gat gct aaa atc tat gat gaa ata tgg caa gct ttt gct gtc 1453
141 Ile Arg Asp Ala Lys Ile Tyr Asp Glu Ile Trp Gln Ala Phe Ala Val
142                               450                               455                               460
144 ttc tta cca gtg aaa act gtt gga gta caa gga gac caa aga acc cat 1501
145 Phe Leu Pro Val Lys Thr Val Gly Val Gln Gly Asp Gln Arg Thr His
146                               465                               470                               475
148 tcc cac gct gtt gca ctt aga gca gtc aca agt caa gat gga atg act 1549
149 Ser His Ala Val Ala Leu Arg Ala Val Thr Ser Gln Asp Gly Met Thr
150 480                               485                               490                               495
152 gca gac tgg tac tac ttt gat ttc aag ttc ctt gac gac gta tca aga 1597
153 Ala Asp Trp Tyr Tyr Phe Asp Phe Lys Phe Leu Asp Asp Val Ser Arg
154                               500                               505                               510
156 aag atc tgc aat agt gtt cgt ggt gta aat cga gtt ctg ctg gat att 1645
157 Lys Ile Cys Asn Ser Val Arg Gly Val Asn Arg Val Leu Leu Asp Ile
158                               515                               520                               525
160 aca tca aag cct cca tca aca atc gaa tgg gaa taatttgta taaagaatgc 1698
161 Thr Ser Lys Pro Pro Ser Thr Ile Glu Trp Glu
162                               530                               535
164 tatatttggt gaccaaagta ggattctttt gtgatttttg gtgcataaca aaaaggaaga 1758
166 aaatcataat agaaatttag gtccttttgt tatgtggtag aactggttct tgggtaatta 1818
168 tgtgcaatgc totcaacaat tttgtatgtt tatgggtatg atgataccaa attttactca 1878
170 gatcttggtg gtacattttt cttatccaag tatagtaaca tgtggccagg catcaaaagc 1938
172 ctattccact caaaaaaaaaa aaaaaaaaaac tcgag 1973
175 <210> SEQ ID NO: 2
176 <211> LENGTH: 538
177 <212> TYPE: PRT
178 <213> ORGANISM: Nicotiana tabacum
180 <400> SEQUENCE: 2
181 Met Glu Pro Gln Thr Gln Ala Lys Lys Ser Asn Leu Val Leu Ile Leu
182 1 5 10 15
184 Asp Tyr Gly Ser Gln Tyr Thr His Leu Ile Thr Arg Arg Ile Arg Ser
185 20 25 30
187 Leu Ser Ile Phe Ser Leu Thr Ile Asn Gly Thr Ser Ser Leu Asp Ser
188 35 40 45
190 Ile Lys Glu Leu Asp Pro Arg Val Ile Ile Leu Ser Gly Gly Pro His
191 50 55 60
193 Ser Val His Ala Asp Gly Ala Pro Cys Phe Pro Pro Gly Phe Ile Glu
194 65 70 75 80
196 Tyr Val Glu Ser Arg Gly Ile His Val Leu Gly Ile Cys Tyr Gly Leu
197 85 90 95
199 Gln Leu Ile Val Gln Lys Leu Gly Gly Val Val Lys Ile Gly Glu Lys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/088,027

DATE: 04/04/2002

TIME: 15:39:20

Input Set : A:\seq list.txt

Output Set: N:\CRF3\04042002\J088027.raw

200				100				105				110				
202	His	Glu	Tyr	Gly	Arg	Met	Glu	Ile	Glu	Val	Gly	Lys	Asn	Val	Val	Gly
203			115					120					125			
205	Gly	Leu	Phe	Gly	Asn	Thr	Glu	Ile	Gly	Asp	Lys	Gln	Val	Val	Trp	Met
206		130					135					140				
208	Ser	His	Gly	Asp	Glu	Ala	Val	Lys	Leu	Pro	Glu	Gly	Phe	Glu	Val	Val
209	145					150					155					160
211	Ala	Arg	Ser	Ser	Gln	Gly	Ala	Val	Ala	Ala	Ile	Glu	Asn	Arg	Glu	Arg
212					165						170					175
214	Arg	Phe	Tyr	Gly	Leu	Gln	Tyr	His	Pro	Glu	Val	Thr	His	Ser	Thr	Glu
215				180					185					190		
217	Gly	Met	Arg	Thr	Leu	Arg	His	Phe	Leu	Phe	Asp	Val	Cys	Gly	Val	Thr
218		195						200					205			
220	Ala	Gly	Trp	Lys	Met	Glu	Asp	Val	Leu	Glu	Glu	Glu	Ile	Lys	Val	Ile
221		210					215					220				
223	Lys	Gly	Met	Val	Gly	Pro	Glu	Asp	His	Val	Ile	Cys	Ala	Leu	Ser	Gly
224	225					230					235					240
226	Gly	Val	Asp	Ser	Thr	Val	Ala	Ala	Lys	Leu	Val	His	Lys	Ala	Ile	Gly
227					245					250						255
229	Asp	Arg	Leu	His	Cys	Val	Phe	Val	Asp	Asn	Gly	Leu	Leu	Arg	Tyr	Lys
230			260						265					270		
232	Glu	Arg	Glu	Arg	Val	Met	Glu	Leu	Phe	Glu	Lys	Arg	Leu	His	Leu	Pro
233		275						280					285			
235	Val	Thr	Cys	Val	Asp	Ala	Thr	Glu	Glu	Phe	Leu	Ser	Lys	Leu	Lys	Gly
236		290					295					300				
238	Val	Thr	Glu	Pro	Glu	Met	Lys	Arg	Lys	Ile	Ile	Gly	Lys	Glu	Phe	Ile
239	305					310					315					320
241	Asn	Ile	Phe	Asp	Leu	Phe	Ala	His	Asp	Val	Glu	Glu	Lys	Val	Gly	Lys
242					325					330						335
244	Lys	Pro	Ser	Tyr	Leu	Val	Gln	Gly	Thr	Leu	Tyr	Pro	Asp	Val	Ile	Glu
245			340						345					350		
247	Ser	Cys	Pro	Pro	Pro	Gly	Ser	Gly	Arg	Thr	His	Ser	His	Thr	Ile	Lys
248		355						360					365			
250	Ser	His	His	Asn	Val	Gly	Gly	Leu	Pro	Lys	Asp	Met	Lys	Leu	Lys	Leu
251		370					375					380				
253	Ile	Glu	Pro	Leu	Lys	Leu	Leu	Phe	Lys	Asp	Glu	Val	Arg	Glu	Leu	Gly
254	385					390					395					400
256	Lys	Ile	Leu	Asp	Ile	Ser	Glu	Asp	Phe	Leu	Lys	Arg	His	Pro	Phe	Pro
257				405						410						415
259	Gly	Pro	Gly	Leu	Ala	Val	Arg	Ile	Pro	Gly	Asp	Val	Thr	Ala	Gly	Asn
260				420					425					430		
262	Ser	Leu	Asp	Ile	Leu	Arg	Gln	Val	Asp	Glu	Ile	Phe	Ile	Gln	Ser	Ile
263		435						440					445			
265	Arg	Asp	Ala	Lys	Ile	Tyr	Asp	Glu	Ile	Trp	Gln	Ala	Phe	Ala	Val	Phe
266		450					455					460				
268	Leu	Pro	Val	Lys	Thr	Val	Gly	Val	Gln	Gly	Asp	Gln	Arg	Thr	His	Ser
269	465					470					475					480
271	His	Ala	Val	Ala	Leu	Arg	Ala	Val	Thr	Ser	Gln	Asp	Gly	Met	Thr	Ala
272				485					490							495

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/088,027

DATE: 04/04/2002

TIME: 15:39:20

Input Set : A:\seq list.txt

Output Set: N:\CRF3\04042002\J088027.raw

```

274 Asp Trp Tyr Tyr Phe Asp Phe Lys Phe Leu Asp Asp Val Ser Arg Lys
275                               500                               505                               510
277 Ile Cys Asn Ser Val Arg Gly Val Asn Arg Val Leu Leu Asp Ile Thr
278                               515                               520                               525
280 Ser Lys Pro Pro Ser Thr Ile Glu Trp Glu
281                               530                               535
284 <210> SEQ ID NO: 3
285 <211> LENGTH: 1232
286 <212> TYPE: DNA
287 <213> ORGANISM: Physcomitrella patens
289 <220> FEATURE:
290 <221> NAME/KEY: CDS
291 <222> LOCATION: (3)..(1148)
293 <400> SEQUENCE: 3
294 ga att cgg cac gag gcc act agt acg cag ggt aat att gcc gct att      47
295   Ile Arg His Glu Ala Thr Ser Thr Gln Gly Asn Ile Ala Ala Ile
296     1                               5                               10                               15
298 gaa aat gtg gat tcc aga atc tac gcc ctc caa tac cat ccc gag gtt      95
299 Glu Asn Val Asp Ser Arg Ile Tyr Ala Leu Gln Tyr His Pro Glu Val
300                               20                               25                               30
302 acg cac tca gag aaa ggg aca gag act ttg aga cac ttt ttc ctg aat      143
303 Thr His Ser Glu Lys Gly Thr Glu Thr Leu Arg His Phe Phe Leu Asn
304                               35                               40                               45
306 gtc tgc ggc atg aag gct gac tgg cag atg cag aat gtg ttg gag gaa      191
307 Val Cys Gly Met Lys Ala Asp Trp Gln Met Gln Asn Val Leu Glu Glu
308                               50                               55                               60
310 gag att aaa aag gtc act gcg acc gtc ggc cca gat gat cat gtt att      239
311 Glu Ile Lys Lys Val Thr Ala Thr Val Gly Pro Asp Asp His Val Ile
312     65                               70                               75
314 tgt gca ctc tcc ggg ggc gtg gac tca aca gta gca gct act ctg gtg      287
315 Cys Ala Leu Ser Gly Gly Val Asp Ser Thr Val Ala Ala Thr Leu Val
316     80                               85                               90                               95
318 cac cgt gct att gga gat cgc ctt cat tgt gtg ttt gta gat aat ggc      335
319 His Arg Ala Ile Gly Asp Arg Leu His Cys Val Phe Val Asp Asn Gly
320                               100                              105                              110
322 ctt tgc aga tac aag gaa aga gaa aga gtg atg gcc aca ttt gtg aaa      383
323 Leu Cys Arg Tyr Lys Glu Arg Glu Arg Val Met Ala Thr Phe Val Lys
324                               115                              120                              125
326 gac ctt cat ctg cca gtc act tgt gtg gat gcc act gag cag ttt ctc      431
327 Asp Leu His Leu Pro Val Thr Cys Val Asp Ala Thr Glu Gln Phe Leu
328                               130                              135                              140
330 agc aaa ttg aag ggc gtg gta gat cca gag aga aag agg aag atc atc      479
331 Ser Lys Leu Lys Gly Val Val Asp Pro Glu Arg Lys Arg Lys Ile Ile
332     145                              150                              155
334 gga gca gag ttt att gca gtc ttt gat gaa ttt tcg cac aga ttg gag      527
335 Gly Ala Glu Phe Ile Ala Val Phe Asp Glu Phe Ser His Arg Leu Glu
336     160                              165                              170                              175
338 aga gag att gga aag atg cct gct ttc ctt gtg cag gga aca ctt tat      575
339 Arg Glu Ile Gly Lys Met Pro Ala Phe Leu Val Gln Gly Thr Leu Tyr

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/088,027

DATE: 04/04/2002

TIME: 15:39:21

Input Set : A:\seq list.txt

Output Set: N:\CRF3\04042002\J088027.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date